Homework due Thursday, Sep 22:
Sec 2.1: 1 (justify your answers), 3, 4, 5, 6, 7, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 25, 35, 38

Start reading chapter 2. Just as Chapter 1 generalized ideas about vectors familiar from math 54, Chapter 2 will generalize ideas about matrices:

Ma54 => Ma110
real numbers => fields
vectors of real numbers => vector spaces
  [ Ex: (x1,x2) => also P_n(F), Func(R,R), etc ]
lines and planes through 0 => subspaces
  [ Ex: a*(x1,x2) => span({s1,s2,...,sn}) ]
lines are 1D, planes 2D... => dimension of any vector space

matrices => linear transformations
  [ Ex: [ 1 2 ] => also differentiation, integration etc ]
  [ 3 4 ]
multiplying matrix*vector => applying a linear transformation to a vector
  [ Ex: [ 1 2 ]*[x1]=[ x1+2*x2] => also T(f(x)) = f'(x) ]
  [ 3 4 ] [x2] [3*x1+4*x2]
multiplying matrix*matrix => composing linear transformations
forming inv(Q)*M*Q => changing the basis of a linear transformation

Def: Let V and W be vector spaces over F. A function T: V -> W is a linear transformation from V to W (or just linear) if for all x,y in V and c in F
(a) T(x+y) = T(x) + T(y)
(b) T(c*x) = c*T(x)

(note: (a) => (b) if F = Q, but in general need both)

Lemma:
T(0_V) = 0_W
T linear if and only if T(c*x+y) = c*T(x) + T(y)
T linear implies T(x-y)=T(x) - T(y)
T linear if and only if T(sum_i c_i*v_i) = sum_i c_i*T(v_i)

Ex: V = R^2, W = R, T((x,y)) = x+2*y
ASK & WAIT: Why is $T$ linear?

Ex: $V = W = \mathbb{R}^2$, $T((x;y)) = (x+y; -3x+2y)$

ASK & WAIT: Why is $T$ linear?

This is also written as the matrix-vector multiplication

$$T([x]) = [1 \ 1] \cdot [x] = [x+y]$$
$$([y]) \quad [-3 \ 2] \cdot [y] \quad [-3x+2y]$$

Indeed, $T: \mathbb{F}^n \to \mathbb{F}^m$ where $T$ is multiplying an $n$-vector by an $m \times n$ matrix to get an $m$-vector is a linear transformation.

In section 2.2, we will systematically write *all* linear transformations from $\mathbb{F}^n$ to $\mathbb{F}^m$ as multiplying an $n$-vector by an $m \times n$ matrix, to get an $m$-vector.

But these are not the only linear transformations.

Ex: $V = \{\text{differentiable functions from } [0,1] \to \mathbb{R}\}$
$W = \{\text{functions from } [0,1] \to \mathbb{R}\}$
$T(f(x)) = f'(x) + \int_0^x t^2f(t) \, dt$

ASK & WAIT: Why is $T$ linear? Is $T: V \to V$?

Ex: $V = W = \mathbb{R}^2$, $T_{\theta}(x,y) = \text{vector } (x,y) \text{ rotated clockwise by } \theta$

ASK & WAIT: Why is $T$ linear?

Def: $V = W$, $T(x) = x$ called identity transformation, written $I_V$ or $I$

Def: $T(x) = 0_W$ called zero transformation

Def: Let $T: V \to W$ be linear. Then the null space of $T$, $N(T)$, is the set of all vectors $v$ in $V$ such that $T(v) = 0_W$.

The range space (or just range) of $T$, $R(T)$, is the set of all vectors $w = T(v)$ for all $v$ in $V$

Thm 1: $N(T)$ and $R(T)$ are subspaces of $V$ and $W$, resp.

Proof: Consider $N(T)$: clearly $0_V$ in $N(T)$, since $T(0_V) = 0_W$.

Since $T$ is linear, $v_1$ and $v_2$ in $N(T)$ implies

$$T(a*v_1 + b*v_2) = T(a*v_1) + T(b*v_2) = a*T(v_1) + b*T(v_2)$$
$$= a*0_V + b*0_V = 0_V,$$

so $N(T)$ is closed under $+$ and $*$, and so a subspace.

Next consider $R(T)$: clearly $0_W$ in $R(T)$, since $T(0_V) = 0_W$.

If $w_1$ and $w_2$ in $R(T)$, then there are $v_1$ and $v_2$ so that $w_1 = T(v_1)$ and $w_2 = T(v_2)$ and so
\[ a^1w_1 + b^1w_2 = a^1T(v_1) + b^1T(v_2) = T(a^1v_1) + T(b^1v_2) = T(a^1v_1 + b^1v_2) \]
is in \( R(T) \) too, so \( R(T) \) is closed under + and *, and so a subspace.

**Ex:** \( V = R^2, W = R, T((x;y)) = x + 2y \)

\(-\)

\[ N(T) = \{(x;y): x + 2y = 0\} = \{f*(2, -1) \text{ for all reals } f\} \]

\[ R(T) = \{w: w=x+2y \text{ for some } x, y\} = R \]

ASK & WAIT: Why is \( R(T) = R \)?

**Ex:** \( V = W = R^2, T((x;y)) = (x+y; -3x+2y) \)

\(-\)

\[ N(T) = \{(x;y): x+y=0 \text{ and } -3x + 2y = 0\} \]
solving these 2 equations in 2 unknowns \( \Rightarrow \) \( x = y = 0 \),
so \( N(T) = \{0_V\} \)

\[ R(T) = \{(w,z): w = x+y \text{ and } z = -3x+2y \text{ for some } (x;y)\} \]
solving these 2 equations in 2 unknowns \( \Rightarrow \) \( y = \frac{1}{5}z + \frac{3}{5}w \)
and \( x = -\frac{1}{5}z + \frac{2}{5}w \), so any \( (w;z) \) in range

**Thm 2:** If \( S = \{v_1, ..., v_n\} \) is a basis for \( V \), then

\[ R(T) = \text{span}(T(v_1), ..., T(v_n)) \]

**Proof:** if \( S \) is a basis all \( v \) in \( V \) are of the form

\[ \sum_i a_i^1v_i \], so all \( w \) in \( R(T) \) are of the form

\[ T(\sum_i a_i^1v_i) = \sum_i T(a_i^1v_i) = \sum_i a_i^1T(v_i) \],
so \( \{T(v_1), ..., T(v_n)\} \) is a spanning set.

ASK & WAIT: Is \( \{T(v_1), ..., T(v_n)\} \) necessarily a basis?

**Ex:** \( V = R^2, W = R^3, T((x;y)) = (x+y; -3x+2y, -x + 4y) \)

To compute \( N(T) \), note \( x+y=0 \) and \(-3x+2y=0 \) implies \( x=y=0 \) from before, so \( N(T) = \{0_V\} \) as before.

\[ R(T) = \text{span}(T((1;0)), T((0;1))) = \text{span}((1;-3;-1),(1;2;4)) \]

**Def:** Let \( T: V \rightarrow W \) be a linear transformation with null space \( N(T) \) and range \( R(T) \).
If \( N(T) \) is finite dimensional, we call its dimension the nullity of \( T \), written "nullity(T)".
If \( R(T) \) is finite dimensional, we call its dimension the rank of \( T \), written "rank(T)".

**Thm 3 (Dimension Theorem):** Let \( T: V \rightarrow W \) be a linear transformation.
If \( V \) is finite dimensional, then \( \dim(V) = \text{rank}(T) + \text{nullity}(T) \)

**Ex:** Just as in chapter 1 where we asked what our theorems about vector spaces meant for \( R^3 \) or \( R^n \), in chapter 2 we can ask what our
theorems about linear transformation mean for matrices, especially simple matrices like diagonal ones.

Consider $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ and $T$ is multiplication by an $m \times n$ matrix. Suppose $T$ is diagonal with 0s and 1s on the diagonal, i.e. only some $T_{ii}$ can be nonzero, such as

$$
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
$$

so $\text{rank}(T) + \text{nullity}(T) = \# \text{nonzero columns of } T + \# \text{zero columns of } T = \# \text{columns of } T = n = \text{dim}(V)$ as claimed by Dimension Theorem.

For simplicity we suppose $T_{11} = T_{22} = \ldots = T_{rr} = 1$, and the rest are 0. So $T$ has $r$ nonzero columns and $n-r$ zero columns. Then it is easy to see

$$
T(x) = T([ x_1 ] = [ x_1 ]
([ \ldots ] [ \ldots ]
([ x_r ] [ x_r ]
([ x_{r+1} ] [ 0 ]
([ \ldots ] [ \ldots ] \ldots \text{there are } m-r \text{ zeros}
([ x_n ] [ 0 ]

where there is one zero in the result vector for every zero row in $T$. So we can see that $R(T)$ is the space spanned by all vectors of the form on the right above, which has dimension $r = \# \text{nonzero columns of } T$, and $N(T)$ is all vectors of the form $[ 0; \ldots; 0; x_{r+1}; \ldots; x_n ]$, i.e. a space of dimension $n-r = \# \text{zero columns of } T$ as desired.

Proof of Dimension Theorem:

Since $N(T)$ is a subspace of the finite dimensional space $V$, it is also finite dimensional, and has a basis $\{v_1, \ldots, v_k\}$. This basis can be extended to a basis of $V$ (Replacement Theorem), call it $\{v_1, \ldots, v_n\}$. We claim $\{T(v_{k+1}), \ldots, T(v_n)\}$ is a basis of $R(T)$. Assuming this for a moment, since it contains $n-k$ vectors, we get

$$
\text{dim}(V) = n = k + (n-k) = \text{dim}(N(T)) + \text{dim}(R(T))
\text{as desired.}
$$

To see $\{T(v_{k+1}), \ldots, T(v_n)\}$ is a basis, we have to show it spans $R(T)$ and is independent. But since $\{v_1, \ldots, v_n\}$ is a basis of $V$,

$$
R(T) = \text{span}(T(v_1), \ldots, T(v_k), T(v_{k+1}), \ldots, T(v_n))
= \text{span}( 0, \ldots, 0, T(v_{k+1}), \ldots, T(v_n))
= \text{span}(T(v_{k+1}), \ldots, T(v_n))
$$

so it spans $R(T)$. To see it is independent, suppose it is not, and seek a contradiction: Write
0_W = \sum_{i=k+1 \text{ to } n} a_i T(v_i) \quad \text{where not all } a_i = 0
= T(\sum_{i=k+1 \text{ to } n} a_i v_i) \quad \text{since } T \text{ is linear}
so \sum_{i=k+1 \text{ to } n} a_i v_i \text{ is a vector in } N(T), \text{i.e.}
\sum_{i=k+1 \text{ to } n} a_i v_i = \sum_{j=1 \text{ to } k} b_j v_j
\quad \text{where at least one coefficient } (a_i \text{ or } b_j) \text{ is nonzero.}
But this contradicts the independence of the basis \{v_1, \ldots, v_n\}.

Natural questions to ask about any function \( T \), not just linear ones, are
(1) Is \( T \) one-to-one, i.e. does \( T(x)=T(y) \) imply \( x=y \)?
(2) Is \( T \) onto, i.e. for all \( w \) in \( W \), is there a \( v \) in \( V \) such that \( w=T(V) \)?

These are important ideas because \( T \) is one-to-one and onto if and only if \( T \) is invertible, i.e. an inverse function \( \text{inv}(T): W \to V \) exists.

These are easy questions to answer for linear \( T \), given the rank and nullity:

Thm: Let \( T: V \to W \) where \( V \) and \( W \) are finite dimensional. Then
(1) \( T \) is one-to-one if and only if \( \text{nullity}(T) = 0 \), i.e. \( N(T) = \{0_V\} \)
(2) \( T \) is onto if and only if \( \text{rank}(T) = \text{dim}(W) \)
In particular, since \( \text{rank} + \text{nullity} = \text{dim}(V) \), \( T \) is invertible if and only if
\( \text{nullity}(T)=0 \) and \( \text{dim}(V)=\text{dim}(W) \).

Ex: Consider \( T: F^n \to F^m \) to be multiplication by an \( m \times n \) matrix,
where \( T \) is diagonal with all \( T_{ii} = 1 \). We ask whether \( T \) is one-to-one
and/or onto. There are 3 cases, depending on \( m \) and \( n \):
(1) \( m < n \). Then \( T(x_1; \ldots; x_n) = (x_1; \ldots; x_m) \). So \( R(T) = F^m \)
and \( T \) is onto. But \( T(0; \ldots; 0; x_{m+1}; \ldots; x_n) \) [\( m \) leading zeros]
\hspace{3cm} = 0_W, \text{so nullity}(T) = n-m > 0, \text{and } T \text{ is not one-to-one}
We will see that no \( T \) can be one-to-one if \( m < n \).
(2) \( m = n \). So \( T \) is onto for the same reason as above and
\( T(x_1; \ldots; x_n) = (x_1; \ldots; x_n) = 0_W \) if and only if \( x=0_V \),
so \( \text{nullity}(T)=0 \) and \( T \) is one-to-one. \( T \) is the
identity function.
(3) \( m > n \). Then \( T(x_1; \ldots; x_n) = (x_1; \ldots; x_n; 0; \ldots; 0) \) [\( m-n \) trailing zeros]
So \( \text{rank}(T) = n < m = \text{rank}(W) \) and \( T \) is not onto.
But \( T(v)=0_W \) only if \( v=0_V \), so \( \text{nullity}(T) = 0 \) and \( T \) is one-to-one.
We will see that no \( T \) can be onto if \( m > n \).

Proof: (1) Suppose \( N(T) = \{0_V\} \). Then
\( T(x)=T(y) \Rightarrow T(x-y)=0 \ldots \text{since } T \text{ is linear} \)
\hspace{3cm} \Rightarrow x-y=0_V \ldots \text{since } x-y \text{ in } N(T) = \{0_V\}
so \( T \) is one-to-one.
Conversely, if \( T \) one-to-one \( \Rightarrow T(v) = 0 \) \( \text{only if } v = 0 \) \( \Rightarrow N(T) = \{0\} \)

(2) \( T \) onto \( \Rightarrow R(T) = W \Rightarrow \text{rank}(T) = \text{dim}(W) \).

Conversely, \( \text{rank}(T) = \text{dim}(R(T)) = \text{dim}(W) \), so since \( R(T) \) is a subspace of \( W \), we must have \( R(T) = W \), i.e. \( T \) is onto.

Corollary: Suppose \( T: V \to W \) and \( \text{dim}(V) = \text{dim}(W) \) is finite. Then the following are equivalent:

1. \( T \) is one-to-one
2. \( T \) is onto
3. \( T \) is invertible
4. \( \text{rank}(T) = \text{dim}(V) \)

Proof: To prove them "equivalent", we need to show that if any one of them is true, then all of them are true. To do this we will prove that

1. \( \iff \) 2.
2. \( (1) \land (2) \iff (3) \)
3. \( (4) \iff (1) \)

(1) \( \iff \) (2) because (1) \( \Rightarrow \) \( \text{nullity}(T) = 0 \) \( \Rightarrow \)

\[ \text{rank}(T) = \text{dim}(V) - \text{nullity}(T) = \text{dim}(V) = \text{dim}(W) \Rightarrow (2); \]

now note that all the implications work in the opposite direction too.

((1) and (2)) \( \iff \) (3) by the definition of invertibility

(1) \( \iff \) \( \text{nullity}(T) = 0 \iff \text{rank}(T) = \text{dim}(V) - \text{nullity}(T) = \text{dim}(V) \)

We need that the dimensions are finite for this to be true, as you will see on homework.

Ex: \( T: P_2(\mathbb{R}) \to \mathbb{R}^3 \) is defined by \( T(a_2x^2 + a_1x + a_0) = (a_2+a_1,a_1+a_0,a_0). \)

Now \( \text{dim}(P_2(\mathbb{R})) = \text{dim}(\mathbb{R}^3) = 3 \), so we can apply the Corollary.

Now \( (a_2+a_1,a_1+a_0,a_0)=(0,0,0) \implies a_0=0 \Rightarrow 0 = a_1+a_0 = a_1 \Rightarrow 0 = a_2+a_1 = a_2 \) so \( T \) is one-to-one, and hence is onto and invertible.

ASK & WAIT: What is its inverse?

The next theorem says that a linear transformation \( T \) is uniquely determined if we know what it does to a basis:

Thm: \( V, W \) be vectors spaces over \( F \), and let \( \{v_1, \ldots, v_n\} \) be a basis for \( V. \)

Given any subset \( \{w_1, \ldots, w_n\} \) of \( n \) vectors from \( W \), there is exactly one linear transformation \( T: V \to W \) such that \( T(v_i) = w_i. \)

Proof: Since \( \{v_1, \ldots, v_n\} \) is a basis, for any \( v \) in \( V \) there is a unique linear combination \( v = \sum_{i=1}^{n} a_i v_i. \) We can then define \( T(v) = \sum_{i=1}^{n} a_i w_i. \) We need to prove 3 things about \( T: \)
(1) \( T \) is linear:

\[
T(\text{vb } + \text{va}) = T(\text{vb}) + T(\text{va}) \text{ because } \\
T(\sum_i b_i*\text{v}_i + \sum_i a_i*\text{v}_i) = T(\sum_i (b_i+a_i)*\text{v}_i) \\
    = \sum_i (b_i+a_i)*w_i \\
    = \sum_i b_i*\text{w}_i + \sum_i a_i*\text{w}_i \\
    = T(\sum_i b_i*\text{v}_i) + T(\sum_i a_i*\text{v}_i)
\]

\[
T(c * \text{vb}) = c*T(\text{vb}) \text{ because } \\
T(c* \sum_i b_i*\text{v}_i) = T(\sum_i (c*b_i)*\text{v}_i) \\
    = \sum_i (c*b_i)*w_i \\
    = c* \sum_i b_i*\text{w}_i \\
    = c* T(\sum_i b_i*\text{v}_i)
\]

(2) \( T(\text{v}_k) = w_k \) because

\[
T(\text{v}_k) = T(\sum_i a_i*\text{v}_i) \text{ where } a_k=1 \text{ and the rest are zero} \\
    = \sum_i a_i*\text{w}_i \text{ where } a_k=1 \text{ and the rest are zero} \\
    = w_k
\]

(3) \( T \) is unique because if \( U: V \rightarrow W \) also satisfies \( U(\text{v}_i) = w_i \) then

for all \( v = \sum_i a_i*\text{v}_i \) we get

\[
U(v) = U(\sum_i a_i*\text{v}_i) \\
    = \sum_i U(a_i*\text{v}_i) \text{ by linearity of } U \\
    = \sum_i a_i*U(\text{v}_i) \text{ by linearity of } U \\
    = \sum_i a_i*\text{w}_i \text{ by def of } U(\text{v}_i) \\
    = \sum_i a_i*T(\text{v}_i) \text{ by def of } T(\text{v}_i) \\
    = \sum_i T(a_i*\text{v}_i) \text{ by linearity of } T \\
    = T(\sum_i a_i*\text{v}_i) \text{ by linearity of } T \\
    = T(\text{v})
\]